

Molecular basis of sugar recognition by collectin-K1 and the effects of mutations associated with 3MC syndrome

Venkatraman Girija *et al.*

RESEARCH ARTICLE

Open Access

Molecular basis of sugar recognition by collectin-K1 and the effects of mutations associated with 3MC syndrome

Umakhanth Venkatraman Girija^{1,2}, Christopher M Furze¹, Alexandre R Gingras^{2,3}, Takayuki Yoshizaki⁴, Katsuki Ohtani⁴, Jamie E Marshall¹, A Katrine Wallis⁵, Wilhelm J Schwaeble¹, Mohammed El-Mezgueldi², Daniel A Mitchell⁶, Peter CE Moody², Nobutaka Wakamiya⁴ and Russell Wallis^{1,2*}

Abstract

Background: Collectin-K1 (CL-K1, or CL-11) is a multifunctional Ca^{2+} -dependent lectin with roles in innate immunity, apoptosis and embryogenesis. It binds to carbohydrates on pathogens to activate the lectin pathway of complement and together with its associated serine protease MASP-3 serves as a guidance cue for neural crest development. High serum levels are associated with disseminated intravascular coagulation, where spontaneous clotting can lead to multiple organ failure. Autosomal mutations in the CL-K1 or MASP-3 genes cause a developmental disorder called 3MC (Carnevale, Mingarelli, Malpuech and Michels) syndrome, characterised by facial, genital, renal and limb abnormalities. One of these mutations (Gly²⁰⁴Ser in the CL-K1 gene) is associated with undetectable levels of protein in the serum of affected individuals.

Results: In this study, we show that CL-K1 primarily targets a subset of high-mannose oligosaccharides present on both self- and non-self structures, and provide the structural basis for its ligand specificity. We also demonstrate that three disease-associated mutations prevent secretion of CL-K1 from mammalian cells, accounting for the protein deficiency observed in patients. Interestingly, none of the mutations prevent folding or oligomerization of recombinant fragments containing the mutations *in vitro*. Instead, they prevent Ca^{2+} binding by the carbohydrate-recognition domains of CL-K1. We propose that failure to bind Ca^{2+} during biosynthesis leads to structural defects that prevent secretion of CL-K1, thus providing a molecular explanation of the genetic disorder.

Conclusions: We have established the sugar specificity of CL-K1 and demonstrated that it targets high-mannose oligosaccharides on self- and non-self structures via an extended binding site which recognises the terminal two mannose residues of the carbohydrate ligand. We have also shown that mutations associated with a rare developmental disorder called 3MC syndrome prevent the secretion of CL-K1, probably as a result of structural defects caused by disruption of Ca^{2+} binding during biosynthesis.

Keywords: Complement activation, structural biology, C-type lectin, 3MC syndrome

Background

Collectin-K1 (CL-K1) is a multifunctional secreted pattern-recognition lectin with key roles in host defence, tissue homeostasis and embryogenesis [1,2]. It binds to pathogen-associated molecular patterns and activates mannan-binding lectin (MBL)-associated serine proteases

(MASPs-1, -2 and -3) to initiate the lectin pathway of complement activation, and stimulate immune and inflammatory processes [3,4]. It also binds to apoptotic cells and may function as a scavenger receptor to facilitate their clearance as has been shown for other collectins. Elevated serum levels are associated with a severe clotting disorder called disseminated intravascular coagulation, characterised by systemic coagulation and microvascular thrombi, that often leads to multiple organ failure [5]. During embryogenesis, CL-K1 is highly expressed in

* Correspondence: rw73@le.ac.uk

¹Department of Infection, Immunity and Inflammation, University of Leicester, Leicester LE1 9HN, UK

²Department of Biochemistry, University of Leicester, Leicester LE1 9HN, UK
Full list of author information is available at the end of the article

craniofacial cartilage, heart, bronchi, kidney, and vertebral bodies and together with MASP-3 probably serves as a guidance cue for neural crest cell migration by recognising specific endogenous carbohydrate epitopes [6]. Separate autosomal mutations in the genes encoding CL-K1 or MASP-3 lead to a rare autosomal recessive disorder called Carnevale, Mingarelli, Malpuech and Michels (3MC) syndrome, associated with craniofacial dysmorphism, mental retardation and genital, renal and limb abnormalities, highlighting the importance of these complexes during development.

CL-K1 is a member of the collectin family of animal lectins with an N-terminal collagen-like domain linked to C-terminal carbohydrate-recognition domains (CRDs) via a helical neck [7]. The collagenous stems are tethered together at the N-terminus to form characteristic bouquet- or fan-shaped structures. CL-K1 can form homooligomers as well as heterooligomers with the related molecule collectin-L1 (CL-L1) [8]. Clusters of CRDs recognise their carbohydrate targets through multiple weak interactions [9]. Ca^{2+} plays an essential role in this process by forming a ternary complex with the protein and the glycan ligand and by stabilising the CRDs [10–12]. CL-K1 binds to a variety of intact bacteria, fungi, and viruses including *Escherichia coli*, *Candida albicans* and Influenza A virus [3]. It also binds to DNA, explaining at least in part how it targets apoptotic cells [13]. Although selective for mannose and fucose, CL-K1 binds only weakly to monosaccharides (IC_{50} approximately 20 mM) compared to MBL (K_D approximately 1 mM) [14] and other collectins, and little is known about its specificity towards oligosaccharides on self- or non-self structures [3].

Three separate disease-associated mutations have been identified in the coding region of the CL-K1 gene. All lead to changes in the primary structure of the CRD: two result in single amino acid substitutions: Ser¹⁶⁹Pro and Gly²⁰⁴Ser and the third leads to the deletion of Ser²¹⁷. CL-K1 is undetectable in the serum of individuals homozygous for the Gly²⁰⁴Ser mutation, implying that it is either not secreted or is degraded in serum. However, the other disease-associated mutations have not been characterised [6].

Here we show that CL-K1 primarily recognises a subset of high-mannose oligosaccharides containing the disaccharide motif: Man(α 1-2)Man found on both self and non-self structures. It binds to both sugar moieties, in contrast to other collectins that only recognise the terminal sugar. We also show that all three mutations associated with 3MC syndrome prevent normal secretion from mammalian cells probably as a result of structural changes caused by the failure to bind Ca^{2+} during biosynthesis. The protein deficiency would prevent the normal recognition processes of CL-K1/MASP-3 complexes during development, leading to the 3MC phenotype.

Results

Sugar specificity of CL-K1 towards self and non-self ligands

To characterise the sugar specificity of CL-K1, we initially screened a selection of glycoproteins with well-characterised glycans (Figure 1A). Blot analysis revealed binding to immunoglobulin M (IgM), thyroglobulin and yeast invertase and mannan, but not to RNase B, fetuin or IgG. The most discernible difference in the pattern of recognition was the presence of high-mannose oligosaccharides on the glycoproteins that were recognised, compared to predominantly complex sugars on those that were not. An exception was RNase B, which possesses a single N-linked glycosylation site, mainly occupied by Man₅ (approximately 60%) but with some Man₆-Man₉ [15]. To explore carbohydrate binding further, the binding kinetics was determined for yeast invertase, mannan and gp120 from HIV that all possess high-mannose oligosaccharides. CL-K1 bound tightly to all three glycoproteins, but not to milk lactoferrin that contains complex sugars (Figure 1B). In each case, the data fitted best to a two-complex model and the kinetic parameters are presented in Table 1. To confirm that carbohydrate binding leads to activation of complement, C3b deposition was measured on mannan in whole human serum depleted of endogenous lectins (Figure 1C). CL-K1 activated complement with only 2.5-fold lower activity than human MBL.

To examine binding to individual glycans, fluorophore-labelled CL-K1 was used to probe a broad screen of mammalian carbohydrates using the Core H Glycan array technology facility at The Consortium for Functional Glycomics [16]. The array consisted of 377 natural and synthetic mammalian glycans attached to the surface of a glass microscope slide via covalent amide linkages. The relative fluorescence signal provides an accurate indication of the order of affinities [17]. Consistent with the glycoprotein data, the best ligands were all high-mannose oligosaccharides (Figure 2A and Additional file 1), but not all of these were recognised. Mammalian high-mannose structures comprise a Man₅ core attached to the protein via two GlcNAc residues, with α 1-2 linked sugars at the non-reducing termini (Figure 2C). CL-K1 only recognised those structures containing at least one terminal Man(α 1-2)Man epitope. For example, it bound tightly to Man₆, Man₇, Man₈ and Man₉ but not to the Man₅ core (ligand 310) that comprises only α 1-3 and α 1-6 linkages. It also bound to smaller glycans containing the essential disaccharide epitope, for example, the trisaccharide, Man(α 1-2)Man(α 1-3)Man (ligand 189). The GlcNAc residues were neither necessary nor enhanced binding (for example, compare Man₉GlcNAc₂, ligand 192, and Man₉, ligand 312). This pattern of recognition is different from other collectins, such as MBL and pulmonary-surfactant protein D that typically recognise the terminal sugar alone (mannose, fucose or N-acetylglucosamine) [18,19]. It is

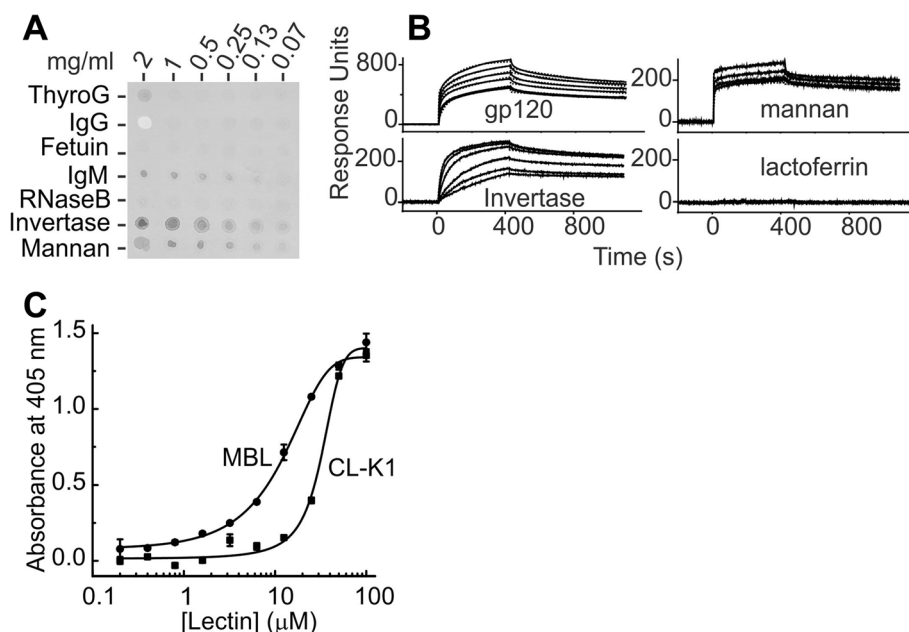


Figure 1 Glycoprotein binding and complement activation by CL-K1. **A**) Immunoblot of CL-K1 binding to immobilised mammalian and yeast glycoproteins. **B**) Surface plasmon resonance of CL-K1 binding to glycoproteins. CL-K1 was immobilised on the chip surface (6,000 response units) and six two-fold dilutions of each glycoprotein were injected at a starting concentration of 200 nM. **C**) Complement component C3b was measured on mannan following incubation with human serum, depleted of endogenous lectins, and supplemented with either CL-K1 or human recombinant MBL. C3b and its breakdown products were detected by absorbance using an anti-human C3c antibody with alkaline phosphatase-conjugated goat anti-rabbit secondary antibody and p-nitrophenyl phosphate as the substrate. Protein concentrations were calculated based on the molecular masses of CL-K1 and MBL trimers (78 kDa in each case). CL-K1, collectin-K1; MBL, mannan-binding lectin.

also different from other C-type lectins, such as DC-SIGN or DC-SIGNR that target larger high-mannose oligosaccharides preferentially (Man₅ to Man₉) [17]. To quantify the differences in affinities, we compared binding to three different epitopes that all form part of a high-mannose oligosaccharide. Inhibition by Man(α1-2)Man (IC₅₀ of 0.02 mM) was 10- and 8-fold greater than for the Man₃ and Man₅ core structures, and >300-fold greater than for mannose alone (Figure 2C). Thus, CL-K1 primarily recognises the terminal disaccharide of the glycan, rather than the terminal sugar alone, explaining its unusual binding properties.

Table 1 Kinetics of binding of CL-K1 to glycoproteins

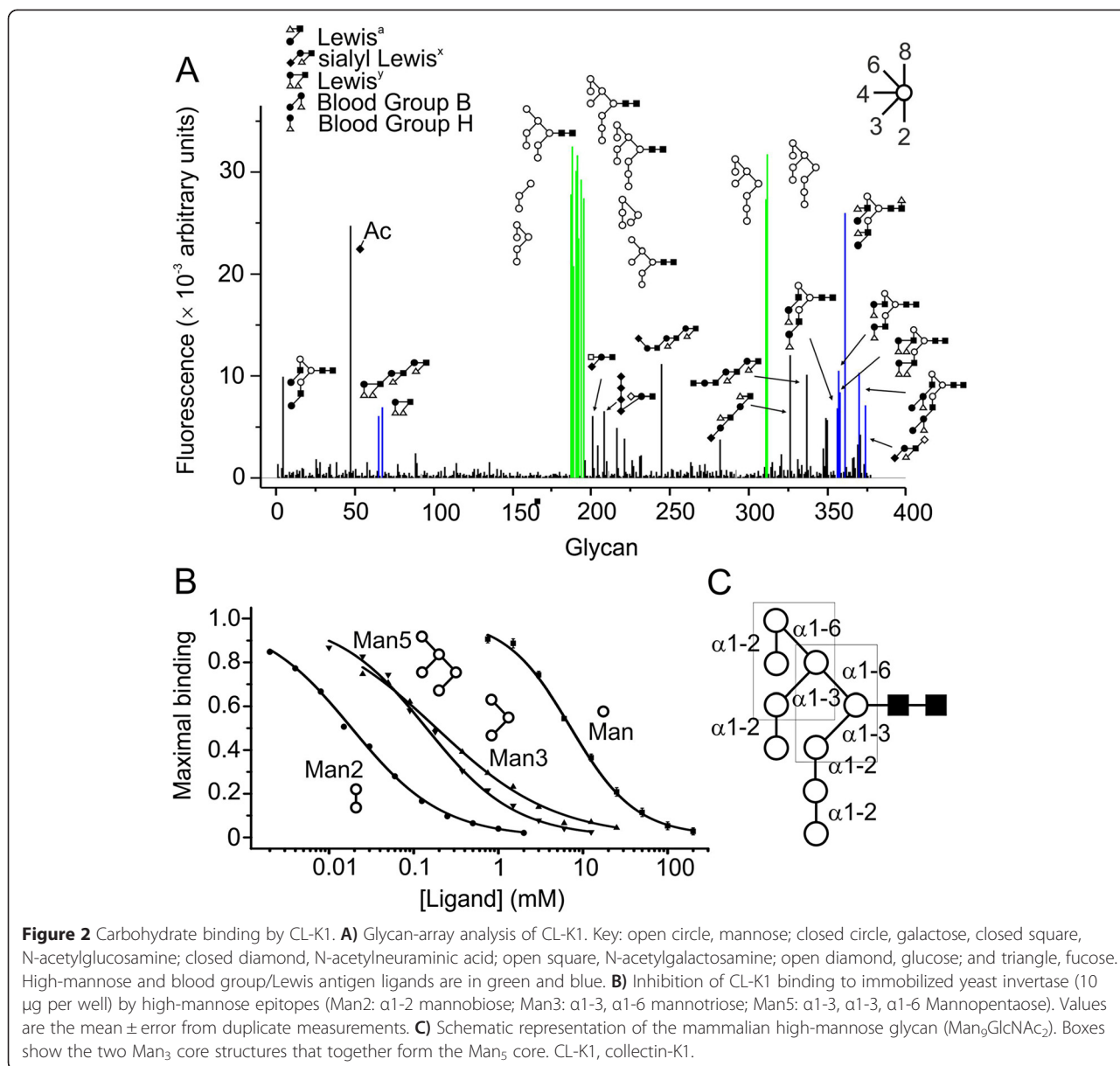
Glycoprotein	k_{on} (M ⁻¹ s ⁻¹)	k_{off} (s ⁻¹)	K_D (nM)
Mannan	$1.6 \pm 0.3 \times 10^3$	$<1.0 \times 10^{-6}$	<0.8
	$1.0 \pm 0.2 \times 10^6$	$3.0 \pm 0.2 \times 10^{-4}$	0.31 ± 0.9
Invertase	$6.8 \pm 0.4 \times 10^4$	$1.2 \pm 0.2 \times 10^{-3}$	18 ± 4
	$6.7 \pm 0.5 \times 10^5$	$<1.0 \times 10^{-6}$	<0.002
gp120	$1.5 \pm 0.3 \times 10^5$	$1.1 \pm 0.2 \times 10^{-3}$	8 ± 4
	$>1 \times 10^6$	$<1.0 \times 10^{-6}$	<0.001

All data were fitted to a two-complex binding model. Values of k_{on} above 1×10^6 M⁻¹ s⁻¹ and k_{off} below 1.0×10^{-6} s⁻¹ exceed the reliable detection limits of the instrument, so are given as lower and upper limits respectively. Data are mean \pm error from two experiments.

CL-K1 also bound to subsets of fucosylated glycans, including those containing Lewis^a, Lewis^y and blood group B antigens as well as to a selection of structures containing blood group H antigens (Figure 2A). Most fucosylated sugars were not recognised, however, suggesting that CL-K1 contacts multiple monosaccharide moieties in each glycan and that the chemical environment of the binding epitope is important. Binding to almost all ligands was Ca²⁺-dependent (see Additional file 1), compatible with the pivotal role of Ca²⁺ in carbohydrate recognition by C-type lectins. An exception was 9-O-acetylneuraminic acid, which gave a similar binding signal in the presence and absence of Ca²⁺, implying that it interacts with a distinct and independent binding site on CL-K1.

Structure of CL-K1 in complex with a disaccharide ligand

In order to determine the precise specificity and mechanism of carbohydrate recognition by CL-K1, a recombinant trimeric fragment comprising the neck and CRDs was produced in *E. coli*, refolded from inclusion bodies and was crystallized in the presence of Ca²⁺, and structures were determined both with and without its primary ligand, Man(α1-2)Man (PDB: 4YLI and 4YMD for the CL-K1 trimer and the trimer bound to Man(α1-2)Man; Table 2). The structures confirm that the CL-K1 fragment is a homotrimer with the three CRDs linked



via an α -helical coiled coil (Figure 3). The asymmetric unit of each crystal contained two trimers with each CRD bound to three Ca^{2+} . Density for the sugar was present for four of the six polypeptides (Figure 3A and B). In the ligand complex, the reducing mannose of the disaccharide binds to the Ca^{2+} common to all C-type lectins, with the equatorial 3- and 4-OHs of the sugar forming coordination bonds with the Ca^{2+} and hydrogen bonds to surrounding Asn, Asp and Glu residues that also serve as Ca^{2+} ligands. The interaction is unusual, however, because it involves the penultimate sugar (internal in a high-mannose oligosaccharide) rather than the terminal sugar (non-reducing) that is recognised by other collectins (Figure 3C and D) [10,19]. In addition, the terminal mannose forms hydrogen bonds to the carboxylate group of Glu²⁴⁴ and the guanidinium group of

Arg²⁰⁰ to form a small hydrogen-bonding network that stabilises the interaction compared to mannose alone, providing an explanation for its binding specificity. Comparison of the bound and unbound structures reveals no major conformational changes upon carbohydrate recognition. The interaction is somewhat reminiscent of glycan binding by DC-SIGN and DC-SIGNR, which also target high-mannose oligosaccharides via internal sugar moieties (Figure 3E). The ligand specificities of CL-K1 and DC-SIGN/DC-SIGNR are significantly different, however, as noted above.

Ca²⁺ stabilises CL-K1 by organising the binding loops of the CRDs

Ca²⁺ is integral to the structure and function of CL-K1 (Figure 3). It not only interacts directly with the

Table 2 Data collection and refinement statistics

Structural data	CL-K1 trimer	CL-K1 + Man ₂
PDB code	4YLI	4YMD
Data collection		
Beamline	Diamond I04	Diamond I04-1
Space group	P 1 2 ₁ 1	P 1 2 ₁ 1
a, b, c, Å	73.0 107.6 77.8	71.9 106.7 78.4
α, β, γ, °	90, 91.5, 90	90, 92.2, 90
Resolution, Å	48.21-2.45 (2.538-2.45)	44.08-2.87 (2.973-2.87)
R _{sym}	0.054 (0.48)	0.052 (0.54)
I/σI	10.1 (2.0)	11.2 (1.8)
Completeness, %	95.5 (87.2)	91.0 (93.6)
Redundancy	3.3 (3.1)	2.9 (2.9)
Refinement		
Resolution, Å	48.21-2.45 (2.538-2.45)	44.08-2.87 (2.973-2.87)
Unique reflections	42224 (3791)	24670 (2518)
R _{work} /R _{free}	0.1869/0.2212	0.2129/0.2607
No of atoms	7332	7220
Protein	7162	7062
Ligands/ion	121	152
water	49	6
B-factors, Å ²	74.5	93.2
macromolecules	74.5	93.0
ligands	76.4	105.6
water	60.7	62.6
rms deviations		
Bond lengths, Å	0.003	0.003
Bond angles, °	0.62	0.77

The highest-resolution shell is shown in parentheses.

carbohydrate ligand but also binds to loops at the top of the CRD. We used the intrinsic fluorescence of the trimeric fragment as a probe to measure Ca²⁺ binding. Each CRD contains two tryptophan residues, which together form part of the hydrophobic core and are sensitive to changes in the conformation of the adjacent Ca²⁺-binding loops. Addition of Ca²⁺ caused a large enhancement of the native fluorescence (approximately 50%) combined with a shift in λ_{em, max} to a shorter wavelength (from 346 nm to 339 nm), as the tryptophans move to a more hydrophobic environment (Figure 4A). The K_D was measured as 0.43 mM by titration with Ca²⁺, with a Hill coefficient of 2.4 (Figure 4B). Thus, although only one Ca²⁺ interacts with the ligand directly, binding itself is highly cooperative, with all three Ca²⁺ being important.

We also investigated the role of Ca²⁺ towards the stability of the CRDs of CL-K1 using urea denaturation. Unfolding of the CRD/neck fragment was accompanied by a dramatic quenching of the fluorescence and a shift in λ_{em, max} to approximately 354 nm as the tryptophans become exposed

to the polar solvent (Figure 4A). Unfolding could be explained by a simple two-state transition, both in the presence and absence of Ca²⁺. However, the midpoint (denaturation concentration) was 2.4 M and 1.4 M urea for the Ca²⁺-bound and -free forms (Figure 4C), indicating that Ca²⁺ stabilises the CRDs.

The three Ca²⁺ interact with residues primarily from loops towards the top of the CRD (Figure 5A). In the absence of Ca²⁺, these regions are likely to be more flexible. To investigate this possibility, we compared the susceptibility of the trimeric CL-K1 fragment towards trypsin, which preferentially cleaves disordered regions of proteins. The Ca²⁺-free form was highly sensitive to trypsin (Figure 4E). By contrast, an approximately 15 kDa fragment was generated in the presence of Ca²⁺ that was relatively resistant to further proteolysis (Figure 4E and F). This fragment was identified as the CRD by Edman degradation, beginning at residue Glu¹⁴⁸. By measuring proteolysis in the presence of increasing concentrations of Ca²⁺, the apparent K_D for calcium was determined as approximately 1 mM, in good agreement with fluorescence data (Figure 4D and F). Thus, Ca²⁺ stabilises the CRD by tethering the binding loops together. In the absence of Ca²⁺, these loops become disordered and more exposed.

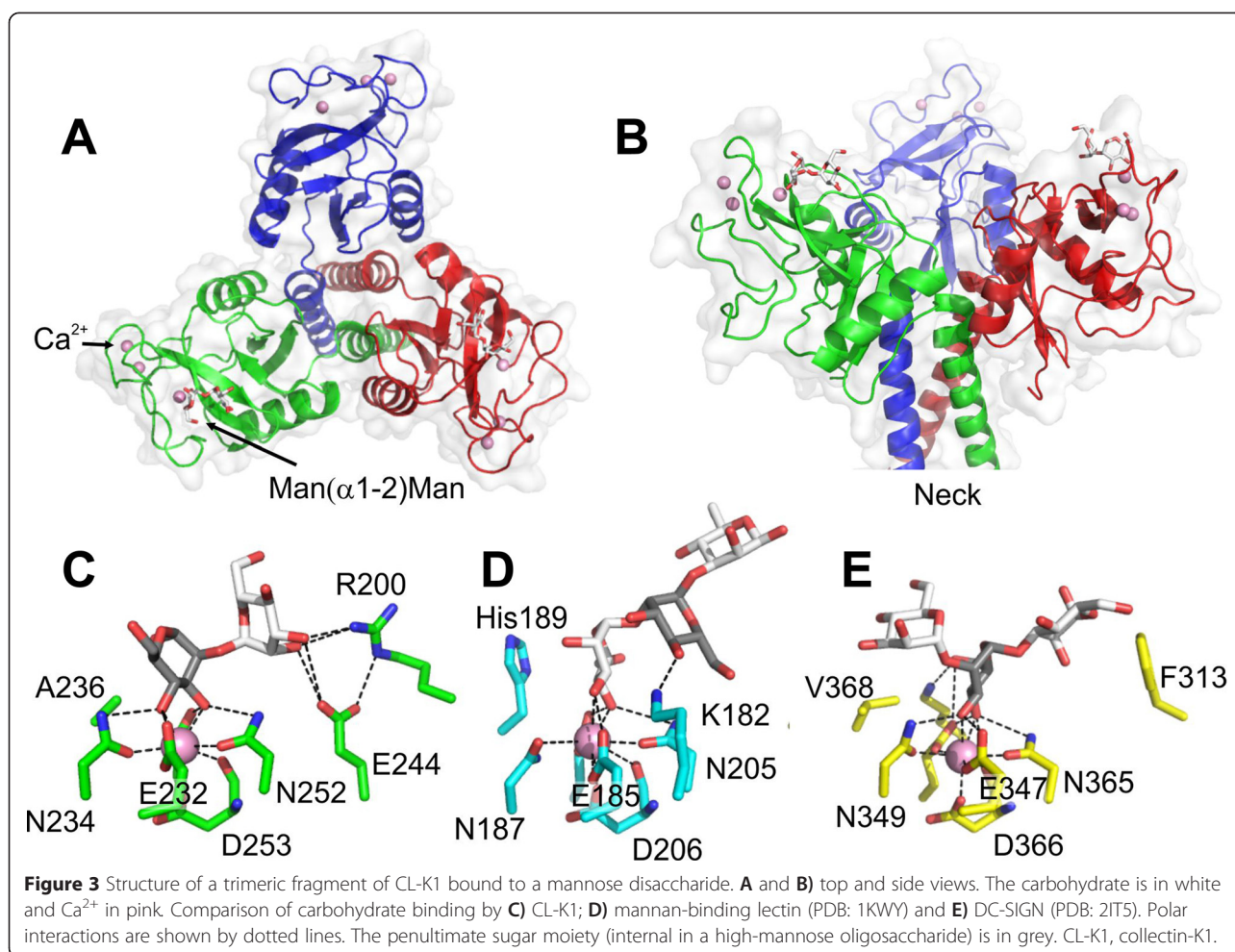
Mutations associated with 3MC syndrome prevent secretion of CL-K1

As well as explaining the mode of carbohydrate recognition, the structures of CL-K1 also reveal the location of residues that are altered in the 3MC syndrome (Ser¹⁶⁹, Gly²⁰⁴ and Ser²¹⁷). Gly²⁰⁴ and Ser²¹⁷ are located at the base of loop L1 (Figure 5A) that forms part of the binding site for two of the three Ca²⁺. The carbonyl group of Gly²⁰⁴ forms a hydrogen bond with the amide of Ser²¹⁷, helping to anchor the loop to the side of the domain. Ser¹⁶⁹ forms part of one of two helices at the base of the CRD, close to the helical neck of the trimer.

To characterise the disease-associated mutations, corresponding changes were introduced into the cDNA of CL-K1, and proteins were expressed in Chinese hamster ovary cells. Immunoblotting of media from transfected cells revealed that although wild-type CL-K1 was secreted as expected, no protein was detected in the media of cells transfected with any of the mutant cDNAs (Figure 5B), implying that the encoded proteins are not secreted. This conclusion would explain the protein deficiency in individuals with the Gly²⁰⁴Ser mutation and suggests that those individuals with Ser¹⁶⁹Pro or ΔSer²¹⁷ mutations are also likely to be deficient in CL-K1.

Defective Ca²⁺ binding by CL-K1 containing disease-associated mutations probably leads to elimination during biosynthesis

Surprisingly, when recombinant fragments of CL-K1 (neck and CRDs) containing the mutations (subsequently



referred to as variant fragments) were produced in *E. coli*, they refolded like the wild-type fragment, and good yields of purified protein were isolated (1 to 2 mg/L of culture). This finding was unexpected because we anticipated that the mutations would severely destabilise the CRDs, thus preventing folding. It was therefore of interest to characterise these fragments in more detail to attempt to understand why the mutations are deleterious and prevent secretion *in vivo*.

Collectins normally fold in the endoplasmic reticulum (ER), in the presence of Ca^{2+} where resident chaperones monitor the folding status of proteins [20]. Misfolded or partially folded proteins with exposed hydrophobic regions are recognised by these chaperones, and are targeted for elimination via the ER-associated protein degradation pathway [21]. Consequently, any structural defects caused by the mutations are likely to prevent secretion. Circular dichroism indicated that all three variants were folded, with spectra characteristic of the α/β structure of the neck + CRD regions (Figure 6B). This finding was confirmed by fluorescence spectroscopy and urea denaturation experiments (Figure 7A and B; described in more

detail below). The first clue that they were defective was from gel filtration analysis. Like many lectins, CL-K1 binds to dextran, so is retarded on gel filtration columns such as Superdex. This binding could be prevented by excess mannose as a competitive inhibitor, or by ethylenediaminetetraacetic acid (EDTA) that chelates the Ca^{2+} (Figure 6A), whereupon the wild-type fragment eluted at the position of a homotrimer (47 kDa). The variant fragments were also trimers, but did not bind to the dextran matrix even in the presence of Ca^{2+} , implying loss of carbohydrate binding. To confirm this possibility, we measured binding to yeast invertase. Both full-length CL-K1 and the wild-type fragment bound to the invertase, but the three variant fragments did not bind (Figure 6C). Thus, although folded and assembled into trimers, the mutations disrupt the CRDs preventing normal interactions with carbohydrate ligands.

The most likely reason for the failure to bind carbohydrate is loss of Ca^{2+} binding. As described above, failure to bind Ca^{2+} destabilises the CRDs and exposes flexible loop regions. The presence of such defects during biosynthesis would likely lead to elimination of CL-K1

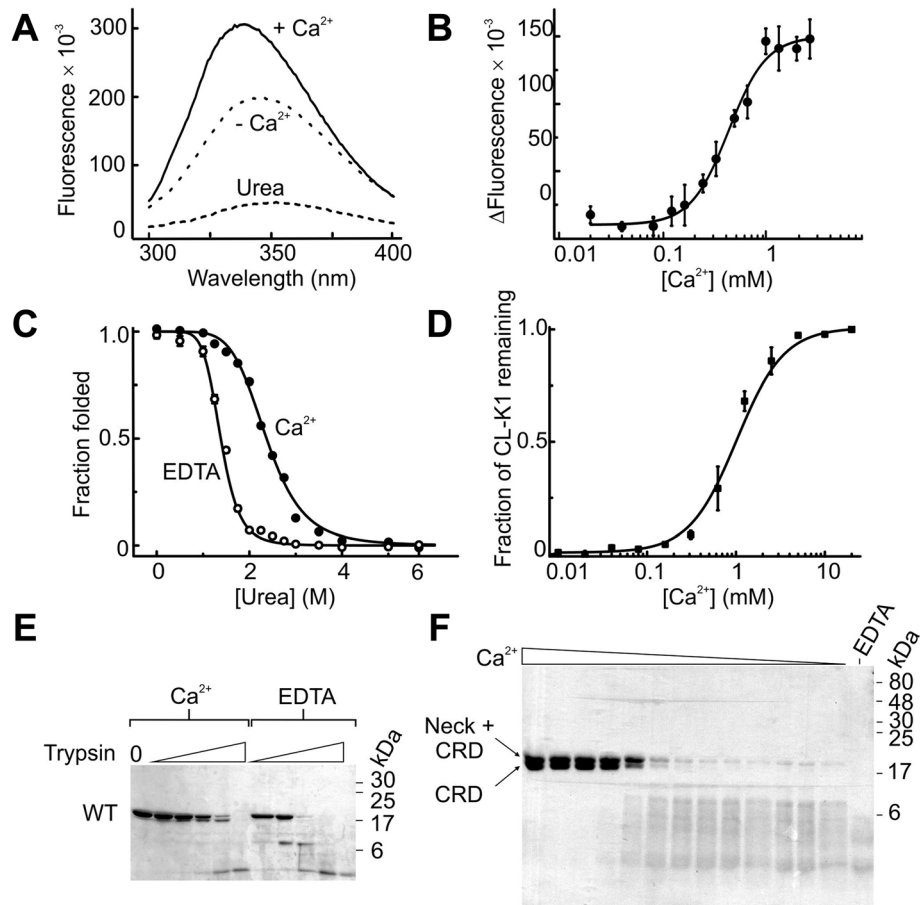


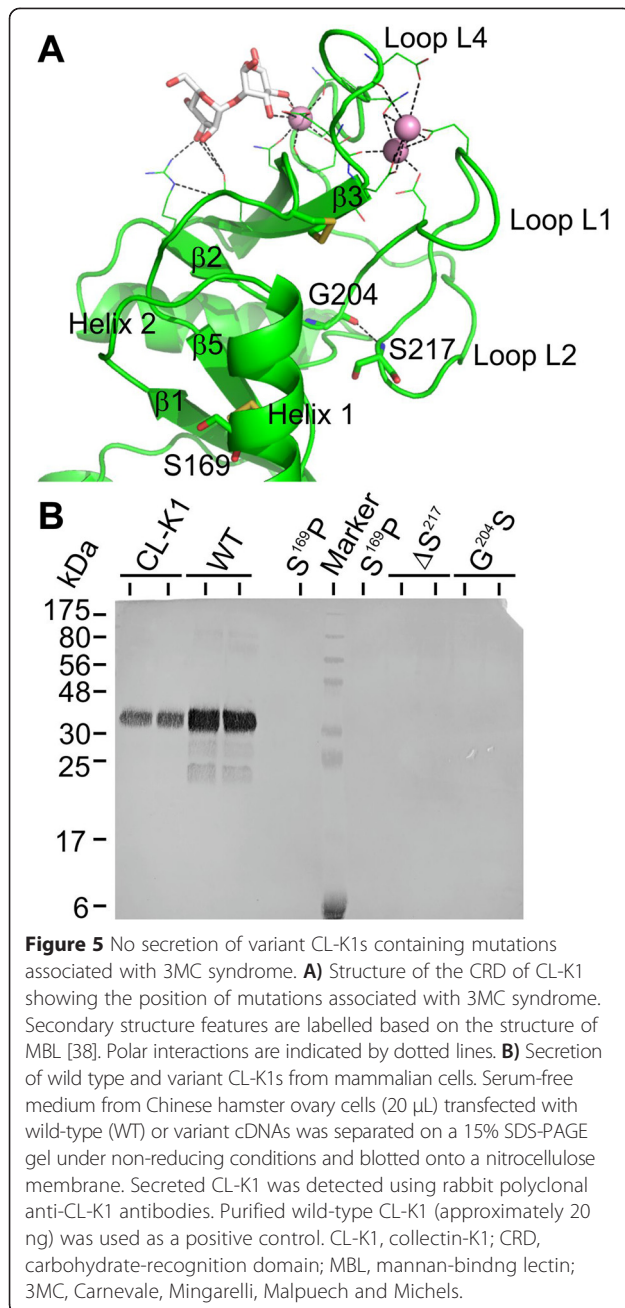
Figure 4 Ca²⁺ stabilises the CRDs of CL-K1. **A**) Fluorescence spectra of the CL-K1 fragment in different buffers. The λ_{ex} was 280 nm. **B**) Ca²⁺ binding to the CL-K1 fragment measured by fluorescence. Measurements were recorded at an λ_{em} of 345 nm. **C**) Urea denaturation in the presence and absence of Ca²⁺ (2 mM). **D**) Proteolysis of CL-K1 in the presence of increasing concentrations of Ca²⁺ (two fold dilutions starting at 5 mM). Samples were separated by SDS-PAGE and gels were scanned to determine the amount of CL-K1 remaining. The amount of trypsin was 0.5% (w/w). The intensities of bands corresponding to the neck and CRD and the CRD alone were combined. A sample gel is shown in **F**). **E**) Proteolysis of CL-K1 in the presence and absence of Ca²⁺. Two fold serial dilutions of trypsin were used at a starting concentration of 2% (w/w). In all experiments data are mean \pm error from two separate experiments. CL-K1, collectin-K1; CRD, carbohydrate-recognition domain.

polypeptides via ER resident chaperones. We therefore tested Ca²⁺ binding using fluorescence assays described above. Addition of Ca²⁺ to the variant CL-K1s did not cause the changes in quantum yield or $\lambda_{em, max}$ that were observed upon Ca²⁺ binding by the wild-type fragment (Figure 7A). Furthermore, urea denaturation curves resembled the Ca²⁺-free form, even in the presence of Ca²⁺, with denaturation concentrations of 1.5, 1.3 and 1.4 M for Ser¹⁶⁹Pro, Gly²⁰⁴Ser, Δ Ser²¹⁷ proteins respectively (Figure 7B). Likewise, each of the variants was equally sensitive to trypsin in the presence or absence of Ca²⁺ (Figure 7B). Thus, all three mutations induce a common structural defect that arises due to the failure to bind to Ca²⁺. *In vivo*, such defects are likely to be detected by chaperones in the ER during biosynthesis leading to elimination. In this way, the mutations would prevent

normal secretion of CL-K1, accounting for the protein deficiency in patients with 3MC syndrome.

Discussion

The data described here explain for the first time how CL-K1 is able to target self- and non-self carbohydrate ligands to fulfil its diverse biological functions. Man(α 1-2)Man epitopes are found at the non-reducing termini of many, but not all, mammalian high-mannose sugars. They are also found on fungal, viral and bacterial structures. For example, mannan from cell walls of *C. albicans*, a known target of CL-K1 [3], is rich in α 1-2 linked mannose oligosaccharides [22]. Likewise, high-mannose oligosaccharides on viral haemagglutinin and gp120 enable CL-K1 to bind to Influenza A virus and HIV [3]. High-mannose oligosaccharides themselves represent only



a small subset of mammalian glycans. Most glycoproteins and glycolipids display complex oligosaccharides that are based on a $\text{Man}_3\text{GlcNAc}_2$ core not targeted by CL-K1, in which the two antennae are decorated with other sugars. The difference in specificity of CL-K1 compared to other collectins is due to the presence of an extended binding site on the CRD (Figure 3C). Whilst unusual for a collectin, additional interactions have been described for other C-type lectins including DC-SIGN, selectins and the macrophage receptor Mincle, each with different ligand preferences [23].

Heterocomplexes of CL-K1 and CL-L1 polypeptides may recognise additional ligands that are not targeted by CL-K1 homooligomers as a result of contacts mediated through the CRD of CL-L1. Like CL-K1, CL-L1 has a preference for ligands possessing mannose and/or fucose epitopes, although little is known about its specificity towards oligosaccharides on self or non-self structures. The presence of CL-K1/CL-L1 complexes as well as CL-K1 homooligomers in serum and in other tissue fluids may serve to increase the range of pathogens targeted by the lectin pathway. It is currently unclear, however, if CL-L1 participates in developmental processes or is involved in the aetiology of 3MC syndrome.

Ca^{2+} is essential for the structure and function of C-type lectins, because it stabilises the CRDs and helps position the loops in the correct arrangement for ligand binding. Although we were unable to crystallise CL-K1 in the absence of Ca^{2+} , the structure of a Ca^{2+} -free form of MBL reveals that the core of the CRD is largely unchanged but the loops are positioned differently with some regions being undetectable in the electron density, implying considerable flexibility [11]. A similar scenario for the CL-K1 variants would explain why the CD spectra are similar to wild type but the CRDs are much more sensitive to proteolysis.

Gly204 and Ser217 are located at the base of Ca^{2+} -binding loop L1 (Figure 5A), so the disease-associated mutations probably interfere with the packing of this loop against the CRD, thereby preventing Ca^{2+} binding. The effects of the Ser¹⁶⁹Pro mutation within Helix 1 are less clear. Nevertheless, the introduction of a proline residue in place of the serine is likely to destabilise the helix and preclude the normal interactions with the CRD. Notably, Tyr163 at the top of the helix forms part of the hydrophobic core and packs against residues in loop 1, forming a hydrogen bond with Asn206. Disruption of these contacts probably prevents Ca^{2+} binding, leading to the disease phenotype.

Folding of CL-K1, like other collectins, probably proceeds in a C- to N-terminal in which folding of the neck and CRDs nucleates assembly of the N-terminal collagen-like domains [20]. All three variant fragments folded and trimerised *in vitro*, so folding is still thermodynamically favourable, despite the mutations. It is thus likely that at least some polypeptides fold *in vivo*. Nevertheless, the CRDs will be defective with exposed, disordered regions, so variants are likely to be retained and subsequently eliminated by the ER-associated protein degradation pathway [21]. The possibility that the mutations prevent folding *in vivo* (but not *in vitro*) cannot be completely ruled out. However, this scenario would seem less likely, in particular because none of the mutations are greatly destabilising *per se* (Figure 7B). Instead, the decrease in stability in each case can be accounted for by the failure to bind to Ca^{2+} .

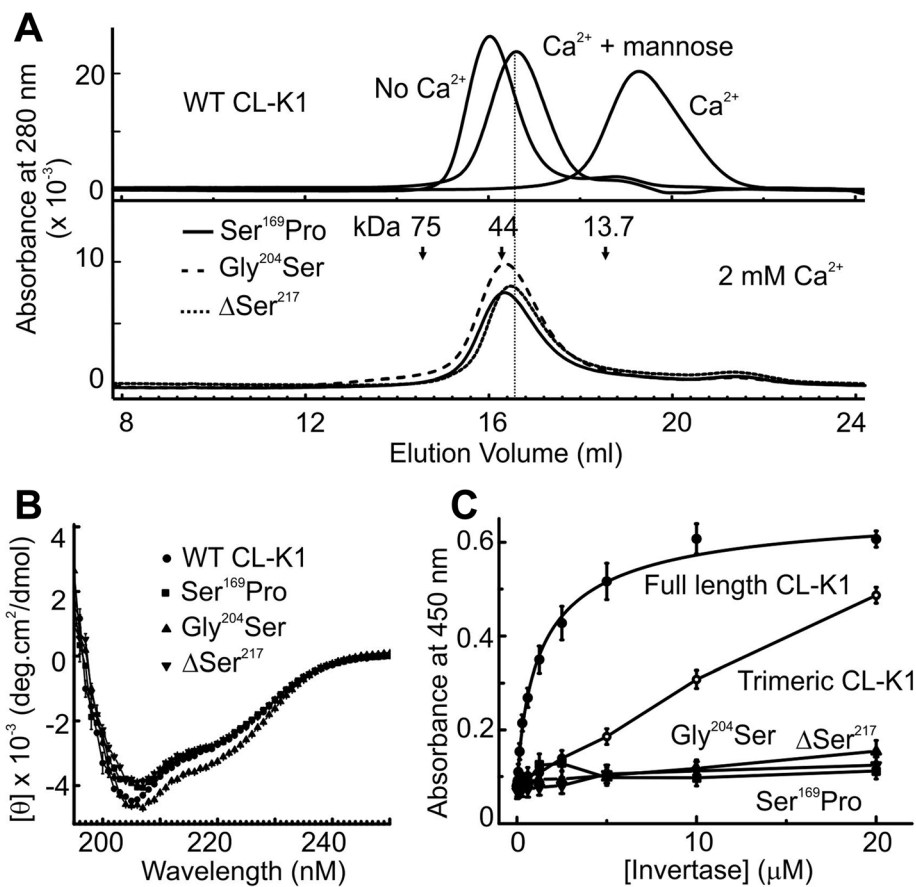


Figure 6 Characterisation of CL-K1 fragments containing disease-associated mutations. **A**) Gel filtration on a Superdex 200 (10/30 column) in 50 mM Tris-HCl at pH 7.4, containing 150 mM NaCl. The elution positions of conalbumin (75 kDa), ovalbumin (44 kDa) and ribonuclease A (13.7 kDa) are indicated. Top, The wild-type CL-K1 fragment interacts with the column matrix in buffer containing Ca²⁺ (2 mM) but not in the absence of Ca²⁺ (2 mM EDTA) or in the presence of inhibitory concentrations of mannose (200 mM). Bottom, None of the three CL-K1 variants interacts with the column, even in the presence of Ca²⁺. **B**) CD spectra of the wild-type CL-K1 fragment and the disease-associated CL-K1 variants. **C**) Binding of biotinylated invertase to full length CL-K1 and trimeric fragments. Values are the mean ± error from duplicate measurements. Tighter binding of the full-length protein is due to its increased avidity. CL-K1, collectin-K1; EDTA, ethylenediaminetetraacetic acid.

The finding that CL-K1 binds to Lewis antigens and certain blood group antigens, including groups B and H, is both interesting and unexpected. Lewis y is expressed mainly during embryogenesis, and in adults expression is restricted to granulocytes and epithelial surfaces [24]. Elevated levels are also found in most epithelial-derived human carcinomas, including breast, ovary, prostate, colon cancers, and high levels are correlated with poor prognosis [24]. Platelets and coagulatory factors such as Von Willebrand Factor are all decorated with blood group antigens, and crosslinking of these components possibly leading to complement activation might account for the link between elevated CL-K1 serum levels and disseminated intravascular coagulation [25]. Notably, activation of the lectin pathway itself promotes activation of the clotting cascade potentially compounding the problem [26]. The larger binding pocket on CL-K1 and the resulting enhanced affinity for its glycan ligands compared to other

collectins suggest that it may be possible to design inhibitors that selectively block ligand binding. These would have therapeutic potential for the treatment of disseminated intravascular coagulation and other disorders associated with profound activation of the innate immune and/or coagulation systems, such as in ischaemia-reperfusion injury.

Conclusions

In conclusion, we have shown that CL-K1 binds to high-mannose oligosaccharides and fucosylated sugars including blood group antigens and Lewis antigens. It binds to high-mannose oligosaccharides via an extended binding site that recognises the terminal two mannose residues. In addition, we have shown that three naturally occurring mutations associated with a severe developmental disorder syndrome prevent secretion of CL-K1 from mammalian cells, probably as a result of a failure to

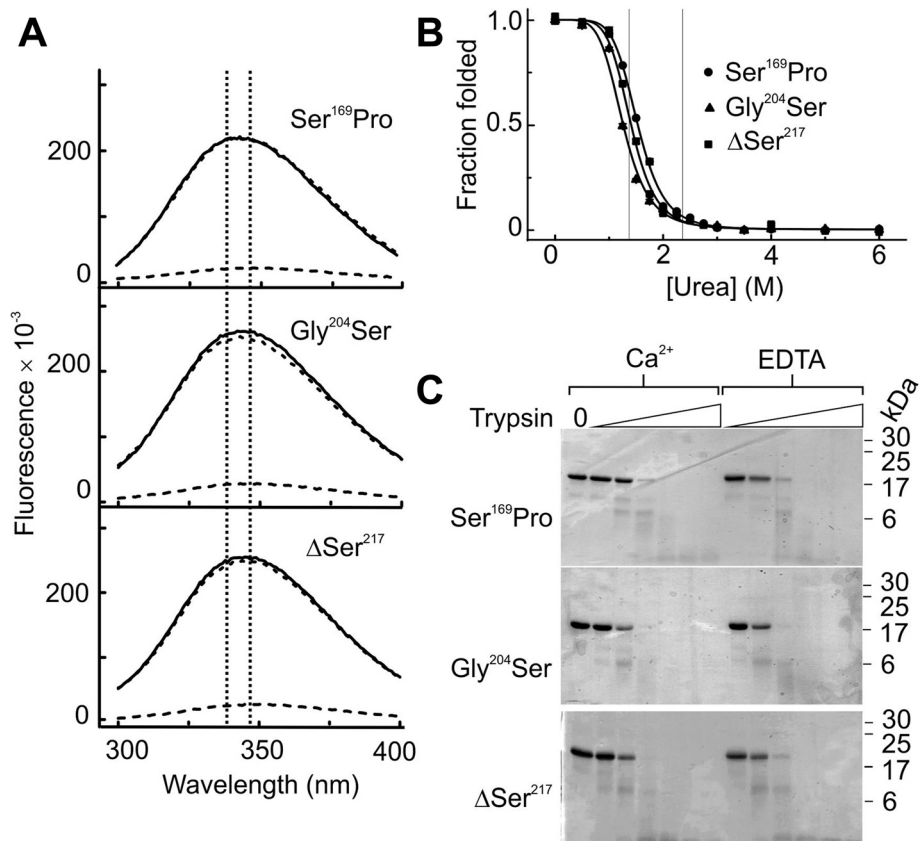


Figure 7 Destabilisation of variant CL-K1s caused by the failure to bind to Ca^{2+} . **A**) Fluorescence spectra of variant CL-K1 trimers in the presence (solid line, 2 mM Ca^{2+}) and absence of Ca^{2+} (dotted line, 2 mM EDTA) and in 6 M urea (dashed line). The $\lambda_{\text{em, max}}$ of Ca^{2+} bound and Ca^{2+} free forms of the wild-type fragment are indicated by vertical dotted lines. **B**) Unfolding of the variant CL-K1 (lower) fragments in urea. All experiments were performed in the presence of 2 mM Ca^{2+} and data were recorded at 345 nm. Values are mean \pm error from duplicate experiments. The denaturation concentrations of Ca^{2+} -bound and Ca^{2+} -free forms of the wild-type fragment are indicated by vertical dotted lines. **C**) SDS-PAGE gels (17.5%) showing proteolysis of variant CL-K1 fragments with two-fold serial dilutions of trypsin (highest concentration is 4% w/w). CL-K1, collectin-K1; EDTA, ethylenediaminetetraacetic acid.

bind Ca^{2+} during biosynthesis. The resulting protein deficiency would prevent normal developmental processes mediated via CL-K1/MASP-3 complexes leading to 3MC syndrome.

Methods

Proteins

Bovine fetuin and RNase B, human milk lactoferrin, yeast invertase and yeast mannan were from Sigma-Aldrich Company Ltd. (Gillingham, UK). Bovine thyroglobulin was from GE Healthcare (Little Chalfont, UK). Human IgM and IgG were from Athens Research & Technology, Inc. (Athens, Georgia, USA). Recombinant HIV gp120 was produced in HEK293 cells [27]. Recombinant human MBL was produced in CHO cells as described for rat MBL [28].

Production of full-length CL-K1 and trimeric fragments

For full-length protein, the cDNA was cloned into the polylinker region of expression vector pED [29] and

protein was produced in Chinese hamster ovary cells and purified by affinity chromatography on mannose-Sepharose columns as described previously for MBL [28,30]. A fragment of the cDNA encoding the neck and CRD (beginning at residue Ser¹¹⁶) was cloned into the polylinker of pET28a. Cells were grown in Power Prime broth (Molecular Dimensions Limited (Newmarket, UK)), induced during mid-log phase with IPTG (1 mM), and harvested after growth at 37°C for an additional 16 hours. Inclusion bodies were isolated and resuspended in 50 mM Tris-HCl, containing 8 M urea and 5 mM dithiothreitol (DTT) and protein was refolded by drop dilution into 50 mM Tris-HCl pH 8.0, containing 9.6 mM NaCl, 0.4 mM KCl, 2 mM MgCl_2 , 2 mM CaCl_2 , 0.5 M arginine, 0.05% polyethylene glycol 3,550, 1 mM GSH and 0.1 mM GSSH at a final protein concentration of 0.1 mg/ml. Proteins were purified by ion exchange chromatography on a 10 mL Q-Sepharose column, using a 0.05 to 1 M gradient of NaCl in 20 mM Tris-HCl at pH 8.0, followed by gel

filtration on a Superdex 75 16/60 column (GE Healthcare) in 20 mM Tris at pH 7.5 containing 50 mM NaCl and 2 mM CaCl₂.

Binding of CL-K1 to glycoproteins

Two-fold serial dilutions of glycoprotein (0.5 µl at a starting concentration of 2 mg/ml) were spotted on to a nitrocellulose membrane and allowed to dry for 30 minutes at room temperature. The membrane was blocked with 5% milk powder in 25 mM Tris-HCl, pH 7.4 containing 150 mM NaCl for two hours, washed with buffer + 0.5% Tween-20 and then incubated with biotinylated CL-K1 (0.06 mg/ml) at room temperature for one hour. After washing, the membrane was incubated with Streptavidin-ALP conjugate (2 µg/ml; Life Technologies Ltd (Paisley, UK)) and developed with BCIP-NBT substrate (Sigma).

Glycan-array screen

Human CL-K1 directly labelled with Alexa-488 was diluted to 200 µg/ml in 20 mM Tris-HCl, pH 7.4 containing 150 mM NaCl, 2 mM MgCl₂, 12 mM CaCl₂, 1% BSA and 0.05% Tween 20. Labelled protein (70 µL) was applied to the printed surface of the microarray chip coated with natural and synthetic glycans *via* amino linkers. The chip was covered and incubated at room temperature in a humidified chamber for one hour. Binding was detected by measuring the fluorescence associated with each glycan. Data are the average of six replicates.

Surface plasmon resonance

CL-K1 was immobilised on the surface of a GLM sensor chip (Bio-Rad Laboratories Ltd (Hemel Hempstead, UK)) by amine coupling. Binding was measured in 10 mM Tris-HCl, pH 7.4 containing 140 mM NaCl, 2 mM CaCl₂ and 0.005% Tween 20 at a flow rate of 25 µl/minute and at 25°C. The protein surface was regenerated in buffer containing 1 M NaCl and 5 mM EDTA.

Complement activation

Complement activation was measured by deposition of complement component C3b on immobilised mannan. Microtiter plates (Nunc) were coated with 2 µg mannan overnight and incubated with 1/100 dilution of whole human serum, previously depleted of endogenous lectins by passage through two mannose-Sepharose columns (2 × 0.5 ml matrix/mL of serum) in 4 mM barbital, 145 mM NaCl, 2 mM CaCl₂, 1 mM MgCl₂, pH 7.4, supplemented with recombinant human MBL or human CL-K1. After incubation at 37°C for one hour and washes with buffer containing 0.5% Tween-20, C3b was detected using rabbit anti-human C3c antibody (Dako UK Ltd (Ely, UK)) with alkaline phosphatase-conjugated goat anti-rabbit secondary antibody (Sigma) and p-nitrophenyl phosphate as the substrate.

Circular dichroism

All measurements were recorded on a Jasco J-715 spectropolarimeter in 10 mM phosphate buffer, pH 7.4 and at 25°C using 6 µM protein.

Fluorescence spectroscopy

All experiments were carried out using a Fluoromax-4 (Horiba Jobin Yvon, Horiba UK Ltd (Middlesex, UK)) instrument with excitation wavelength at 280 nm, slit widths of 2 nm and emission wavelength range of 300 to 400 nm. The λ_{ex} was 280 nm. Protein samples (0.5 µM) in 10 mM Tris-HCl, pH 7.4 containing 140 mM NaCl were measured in the presence of 2 mM CaCl₂, 1 mM EDTA and 6 M urea. For urea denaturation, proteins (0.5 µM) were equilibrated in urea (0 to 6 M) for 15 minutes at room temperature and measurements were recorded at λ_{em} of 345 nm.

Proteolysis

CL-K1 (0.5 mg/ml) was incubated with trypsin (maximum concentration of 4% w/w) in 50 mM Tris-HCl pH 7.5, containing 140 mM NaCl for one hour at 37°C.

Binding assays

MaxiSorb microtiter plates (Nunc) were coated with yeast invertase by overnight incubation in 0.2 M carbonate-bicarbonate buffer at pH 9.6 and at 4°C. Wells were blocked with 5% BSA in 10 mM Tris-Cl, 140 mM NaCl, pH 7.4 and incubated with biotinylated CL-K1 (50 nM) mixed with dilutions of carbohydrate ligand for one hour at room temperature. After washing, wells were incubated with 1:2000 dilution of goat peroxidase conjugated anti-biotin antibody (Sigma) and developed with tetramethylbenzidine (Sigma). Absorbance was measured at 450 nm. CL-K1 binding to invertase was carried out in the same way except that plates were coated with CL-K1 and incubated with dilutions of biotinylated invertase.

Crystallization and structure determination

All crystals were grown using the sitting-drop vapour diffusion method by mixing equal volumes (1.2 + 1.2 µL) of protein and reservoir solution. Protein at approximately 3 mg/ml was mixed with 10% (vol/vol) ethanol in 100 mM HEPES pH 7.0 containing 4 mM CaCl₂. Similar conditions were used to crystallize the complex, except that Man(α1-2)Man (1 mM) was included in the mixture. All crystals were transferred to reservoir solution containing 20% (vol/vol) glycerol, before storage in liquid nitrogen and were maintained at 100 K during data collection. Ligand (1 mM) was included in the cryoprotection buffer for the complex. Diffraction data were collected at beamlines I04 and I04-1 at Diamond Light Source and were processed with iMosflm [31]. Phases were determined by molecular replacement with

Phaser [32] using the structure of pulmonary-surfactant protein D [33] as a search model (PDB: 1B08). Models were optimized by using cycles of manual refinement with Coot [34] and refinement in Refmac5 [35], part of the CCP4 software suite [36], and in Phenix [37].

Additional file

Additional file 1: Glycan-array analysis of CL-K1 in the presence and absence of Ca²⁺.

Abbreviations

3MC: Carnevale, Mingarelli, Malpuech and Michels; BSA: bovine serum albumin; CL-K1: collectin-K1; CRD: carbohydrate-recognition domain; EDTA: ethylenediaminetetraacetic acid; ER: endoplasmic reticulum; MASPs: MBL-associated serine proteases; MBL: mannan-binding lectin.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

UVG, CMF, ARG, TY, KO, JEM, DAM, NW, ME-M, AKW, PCEM and RW carried out the experiments and/or helped to analyse and interpret the data. RW, AKW, DAM and NW wrote the manuscript. UVG, CMF, WJS, RW, NW, PCEM, DAM and ME-M devised the experiments. All authors read and approved the final manuscript.

Authors' information

Umakhanth Venkatraman Girija and Christopher M Furze have Joint first authorship.

Acknowledgements

We thank Diamond Light Source (DLS) for access to beamlines I04 and I04-1, the UK Midlands Block Allocation Group MX8359, and DLS beamline scientists for their help with data collection. We also thank David Smith at Emory University for overseeing the glycan-array screening and Chris Scanlan of the Glycobiology Institute, Oxford for the gift of recombinant HIV gp120. Funding for this work was provided by Medical Research Council UK Grant G1000191/1 (to R.W., P.C.E.M. and W.J.S.). ARG is supported by the American Heart Association Scientist Development Grant (12SDG11610043). NW is supported by the Japanese Grants-in-Aid for Scientific Research (26293124). Resources provided by the Consortium for Functional Glycomics are supported by NIGMS Grant GM62116.

Author details

¹Department of Infection, Immunity and Inflammation, University of Leicester, Leicester LE1 9HN, UK. ²Department of Biochemistry, University of Leicester, Leicester LE1 9HN, UK. ³Department of Medicine, University of California San Diego, La Jolla, CA 92093-0726, USA. ⁴Department of Microbiology and Immunochimistry, Asahikawa Medical University, 2-1-1 Midorigaoka-Higashi, Asahikawa 078-8510, Japan. ⁵Department of Applied Science and Health, Coventry University, Coventry CV1 5FB, UK. ⁶Clinical Sciences Research Laboratories, Warwick Medical School, University Hospital Coventry & Warwickshire Coventry, Coventry CV2 2DX, UK.

Received: 8 January 2015 Accepted: 1 April 2015

References

- Keshi H, Sakamoto T, Kawai T, Ohtani K, Katoh T, Jang SJ, et al. Identification and characterization of a novel human collectin CL-K1. *Microbiol Immunol.* 2006;50:1001–13.
- Motomura W, Yoshizaki T, Ohtani K, Okumura T, Fukuda M, Fukuzawa J, et al. Immunolocalization of a novel collectin CL-K1 in murine tissues. *J Histochem Cytochem.* 2008;56:243–52.
- Hansen S, Selman L, Palaniyar N, Ziegler K, Brandt J, Kliem A, et al. Collectin 11 (CL-11, CL-K1) is a MASP-1/3-associated plasma collectin with microbial-binding activity. *J Immunol.* 2010;185:6096–104.
- Ma YJ, Skjoedt MO, Garred P. Collectin-11/MASP complex formation triggers activation of the lectin complement pathway—the fifth lectin pathway initiation complex. *J Innate Immun.* 2013;5:242–50.
- Takahashi K, Ohtani K, Larvie M, Moyo P, Chigweshe L, Van Cott EM, et al. Elevated plasma CL-K1 level is associated with a risk of developing disseminated intravascular coagulation (DIC). *J Thromb Thrombolysis.* 2014;38:331–8.
- Rooryck C, Diaz-Font A, Osborn DP, Chabchoub E, Hernandez-Hernandez V, Shamseldin H, et al. Mutations in lectin complement pathway genes COLEC11 and MASP1 cause 3MC syndrome. *Nat Genet.* 2011;43:197–203.
- Reid KBM, Colomb MG, Loos M. Complement component C1 and the collectins: parallels between routes of acquired and innate immunity. *Immunol Today.* 1998;12:56–9.
- Henriksen ML, Brandt J, Andrieu JP, Nielsen C, Jensen PH, Holmskov U, et al. Heteromeric complexes of native collectin kidney 1 and collectin liver 1 are found in the circulation with MASPs and activate the complement system. *J Immunol.* 2013;191:6117–27.
- Weis WI, Drickamer K. Trimeric structure of a C-type mannose-binding protein. *Structure.* 1994;2:1227–40.
- Weis WI, Drickamer K, Hendrickson WA. Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature.* 1992;360:127–34.
- Ng KK, Park-Snyder S, Weis WI. Ca²⁺-Dependent structural changes in C-type mannose-binding proteins. *Biochemistry.* 1998;37:17965–76.
- Ng KK, Weis WI. Coupling of prolyl peptide bond isomerization and Ca²⁺ binding in a C-type mannose-binding protein. *Biochemistry.* 1998;37:17977–89.
- Henriksen ML, Brandt J, Iyer SS, Thielens NM, Hansen S. Characterization of the interaction between collectin 11 (CL-11, CL-K1) and nucleic acids. *Mol Immunol.* 2013;56:757–67.
- Iobst ST, Wormald MR, Weis WI, Dwek RA, Drickamer K. Binding of sugar ligands to Ca²⁺-dependent animal lectins. I. Analysis of mannose binding by site-directed mutagenesis and NMR. *J Biol Chem.* 1994;269:15505–11.
- Fu D, Chen L, O'Neill RA. A detailed structural characterization of ribonuclease B oligosaccharides by 1H NMR spectroscopy and mass spectrometry. *Carbohydr Res.* 1994;261:173–86.
- Consortium for Functional Glycomics, functional glycomics gateway. www.functionalglycomics.org.
- Feinberg H, Mitchell DA, Drickamer K, Weis WI. Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science.* 2001;294:2163–6.
- Weis WI, Drickamer K. Structural basis of lectin-carbohydrate interaction. *Annu Rev Biochem.* 1996;65:441–73.
- Shrive AK, Martin C, Burns I, Paterson JM, Martin JD, Townsend JP, et al. Structural characterisation of ligand-binding determinants in human lung surfactant protein D: influence of Asp325. *J Mol Biol.* 2009;394:776–88.
- Heise CT, Nicholls JR, Leamy CE, Wallis R. Impaired secretion of rat mannose-binding protein resulting from mutations in the collagen-like domain. *J Immunol.* 2000;165:1403–9.
- Ruggiano A, Foresti O, Carvalho P. Quality control: ER-associated degradation: protein quality control and beyond. *J Cell Biol.* 2014;204:869–79.
- Johnson MA, Bundle DR. Designing a new antifungal glycoconjugate vaccine. *Chem Soc Rev.* 2013;42:4327–44.
- Taylor ME, Drickamer K. Convergent and divergent mechanisms of sugar recognition across kingdoms. *Curr Opin Struct Biol.* 2014;28C:14–22.
- Madjid Z, Parsons T, Watson NF, Spendlove I, Ellis I, Durrant LG. High expression of Lewis y/b antigens is associated with decreased survival in lymph node negative breast carcinomas. *Breast Cancer Res.* 2005;7:780–7.
- Franchini M, Mannucci PM. ABO blood group and thrombotic vascular disease. *Thromb Haemost.* 2014;112:1103–9.
- Krurup A, Wallis R, Presanis JS, Gal P, Sim RB. Simultaneous activation of complement and coagulation by MBL-associated serine protease 2. *PLoS One.* 2007;2, e623.
- Dunlop DC, Bonomelli C, Mansab F, Vasiljevic S, Doores KJ, Wormald MR, et al. Polysaccharide mimicry of the epitope of the broadly neutralizing anti-HIV antibody, 2G12, induces enhanced antibody responses to self oligomannose glycans. *Glycobiology.* 2010;20:812–23.
- Wallis R, Drickamer K. Molecular determinants of oligomer formation and complement fixation in mannose-binding proteins. *J Biol Chem.* 1999;274:3580–9.
- Kaufman RJ, Davies MV, Wasley LC, Michnick D. Improved vectors for stable expression of foreign genes in mammalian cells by use of the untranslated leader sequence from EMC virus. *Nucleic Acids Res.* 1991;19:4485–90.

30. Wallis R, Drickamer K. Asymmetry adjacent to the collagen-like domain in rat liver mannose-binding protein. *Biochem J.* 1997;325:391–400.
31. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr.* 2011;67:271–81.
32. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr.* 2007;40:658–74.
33. Hakansson K, Lim NK, Hoppe HJ, Reid KBM. Crystal structure of the trimeric α -helical coiled-coil and the three lectin domains of human lung surfactant protein D. *Structure.* 1999;7:255–64.
34. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004;60:2126–32.
35. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr.* 1997;53:240–55.
36. Collaborative Computational Project N. The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst D Biol Crystallogr.* 1994;50:760–3.
37. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* 2010;66:213–21.
38. Weis WI, Kahn R, Fourme R, Drickamer K, Hendrickson WA. Structure of the calcium-dependent lectin domain from a rat mannose-binding protein determined by MAD phasing. *Science.* 1991;254:1608–15.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

